DISCUSSION PAPER SERIES

IZA DP No. 8395

Asymmetry of Information within Family Networks

Joachim De Weerdt
Garance Genicot
Alice Mesnard

August 2014

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

Asymmetry of Information within Family Networks

Joachim De Weerdt
EDI, Tanzania

Garance Genicot
Georgetown University and IZA

Alice Mesnard
City University, London and Institute for Fiscal Studies

Discussion Paper No. 8395
August 2014

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

IZA Discussion Paper No. 8395
August 2014

ABSTRACT

Asymmetry of Information within Family Networks*

This paper studies asymmetry of information and transfers within a unique data set of 712 extended family networks from Tanzania. Using cross-reports on asset holdings, we construct measures of misperception of income among all pairs of households belonging to the same network. We show that there is significant asymmetry of information and no evidence of major systematic over-evaluation or under-evaluation of income in our data, although there is a slight over-evaluation on the part of migrants regarding non-migrants. We develop a static model of asymmetric information that contrasts altruism, pressure and exchange as motives to transfer. The model makes predictions about the correlations between misperceptions and transfers under these competing explanations. Testing these predictions in the data gives support to the model of transfers under pressure or an exchange motive with the recipient holding all the bargaining power.

JEL Classification: O12, O15, D12

Keywords: asymmetric information, transfers, pressure, exchange, altruism

Corresponding author:

Garance Genicot
Department of Economics
Georgetown University
559 ICC, 37th & O Streets, NW
Washington, DC 20057
USA
E-mail: gg58@georgetown.edu

* We thank Stefan Dercon for valuable inputs and comments.

1 Introduction

Private transfers among family members are pervasive in developing countries (Cox and Fafchamps (2008)). For several decades a strand of the literature has been the determinants of these transfers (Lucas and Stark (1985), Rapoport and Docquier (2006)). That literature has, however, typically assumed perfect information between family members.

In contrast, a growing literature on migration suggests that there could be substantial asymmetry of information among close relatives. McKenzie, Gibson, and Stillman (2012) show that Tongo in potential migrants underestimate their future earnings as migrants in New Zealand, and especially so if they have relatives there. Seshan (2013) studies a sample of Indian wives with migrant husbands in Qatar and finds substantial asymmetry of information. Among a small sample of Senegalese migrants in France and Italy, Serror (2012) shows that migrants make mistakes on the asset ownership of the family members remaining in Senegal.

One expects asymmetry of information to be important for private transfers. Using lab experiments in the field to vary the observability of gains of family members, Jakiela and Ozier (2012), in the context of close relatives, and Ambler (2013), in the context of migrants, show that less information decreases transfers to other family members. Seshan (2013) finds that the more wives of Indian migrants working in Qatar under-estimate the overseas income of their husbands, the lower the fraction of income sent home as annual remittances. Batista and Narciso (2013) observe higher remittances from a sample migrants in Ireland when they are offered free phone cards and argue that it is due to improved information.

This paper uses a unique dataset of 712 extended family networks originating the Kagera region in Tanzania to study asymmetry of information and private transfers within these networks. Making use of cross-reports on asset ownership among geographically dispersed extended family networks, we propose a method of measuring of misperceptions of income for each pair of households belonging to the same network. We believe that by using cross-reports on assets our measures will suffer less from measurement error, compared to, for example, directly asking about income; particularly in a context where home consumption plays an important role. However, it raises the question of how to translate these questions on various assets into a household's beliefs regarding the living standards of others. Our method consists of using a weighted sum of differences between believed and actual asset holdings. The weights are set depending on how a particular asset correlates with consumption and allows for negative weights (assets predicting lower consumption) and interaction effects across assets or between assets and household characteristics, such as location. Naturally, since this measure of misperception is based on assets, it will capture asymmetry of information regarding medium term income, but would not capture asymmetry of information regarding short term shocks.

Using this measure we first characterize asymmetry of information among extended family networks. We find substantial levels of asymmetry of information and show that it correlates positively with genetic, social and physical distance between households. Perceptions are, on average, roughly correct. There is also no evidence of large, systematic under- or over-estimates across different types of households. Non-migrants underestimate migrants' income by about 1.8% on average and migrants overestimate non-migrants' income by about 5.2%, which is consistent with McKenzie, Gibson, and Stillman (2012) and Serror (2012). For the median migrant-stayer pair, however, there is no misperception in either direction.

Second, to relate transfers to misperceptions, we develop simple theoretical models of transfers and asymmetry of information.[1] We compare three possible motives for transfers in this model: altruism, in which the potential donor cares about the recipient; pressure, in which the recipient has some means of imposing a utility cost on the donor; and exchange, in which the transfer represents a payment for some good or service provided by the recipient. Our model of altruism predicts a negative partial correlation of transfers with both the recipient's actual income *and* the donor's misperception of that income. Our model of exchange also predicts a negative partial correlation between the transfer and the recipient's income if the donor has all the bargaining power. In contrast, if pressure is the driving motivation or an exchange motive in which the recipient has all the bargaining power, it is the donor's income and the recipient's misperception of the donor's income that matter and are positively correlated with the transfers. Note that these models capture not only the effect of exogenous information on transfers, but also the feedback mechanisms whereby transfers themselves, or the amount requested are informative.

Taking these predictions to the data we determine partial correlation coefficients between transfers and misperceptions. We find that transfers co-move with the recipient's misperception of the donor's wealth, but not with the donor's misperception of the recipient's wealth. This suggests

[1]Since our measure of misperception concerns medium term income and we have one-time observations, the models are static and considers a potential donor and recipient having each some private information regarding their income. This differs from Cole and Kocherlakota (1999), Attanasio and Pavoni (2011) and Hauser and Hopenhayn (2008) who explore the interesting dynamic effects arising from asymmetry of information regarding short term shocks in infinitely-repeated games.

the recipient takes a key place in the relationship, either through exerting pressure to give on the donor in a pressure model or by holding the bargaining power during the exchange of services with the donor. This is consistent with the experimental evidence of Jakiela and Ozier (2012) and Ambler (2013), and the large ethnographic literature highlighting the importance of disapproval, shaming, ostracism and other means of pressure, as described and cited in Platteau (2012) and Chort, Gubert, and Senne (2012).

The rest of the paper is organized as follows. Section 2 presents the data on extended family networks in Tanzania that we use to build the measures of asymmetry of information. These measures are defined and validated in Section 3. Section 4 describes misperceptions of income and how they relate to measures of closeness. Section 5 presents competing models of transfers and how they relate to misperceptions of income between two households belonging to a family network. Section 6 studies the correlations between transfers and misperceptions of income to examine the empirical validity of the models and Section 7 concludes.

2 Data

This paper uses two waves of the Kagera Health and Development Survey (KHDS). The original survey (KHDS 1991-94) interviewed 915 households in 52 villages representative of the Kagera Region of Tanzania over four rounds from 1991 to 1994. This region, in the north-western part of the country, has a population of 2.5 million people, the vast majority of whom depend on agriculture as the main source of income. The Kagera region is relatively isolated: it borders landlocked countries Uganda, Rwanda and Burundi and is 1,400 km away from the main port and commercial capital of the country, Dar es Salaam. Research using the data that underlie this paper has shown that diversification of income generating activities and migration are key household strategies for growth (De Weerdt (2010) and Beegle, De Weerdt, and Dercon (2011), respectively).

KHDS 2010 attempted to trace all individuals on any original household roster and administer a full household interview in the household in which they were found residing in 2010. The survey attained very high recontact rates. Out of the original 915 households there are only 71 households (8%) where not a single individual was traced (excluding 26 households where all members had died). The interviewing team accounted, in 2010, for 88% of the 6,353 individuals listed on any KHDS91-94 roster: 68% of the original respondents were visited and the household in which they lived were administered a household interview, while 20% of respondents were confirmed to have died and the circumstances of their death were recorded through an interview with an informant (often during the household interview with other surviving household members). 12% of individuals were not found.[2] Out of the interviewed individuals 45% were found residing in the baseline village, 53 percent had migrated within the country, 2 percent to another East African country (primarily Uganda) and 0.3 percent had moved outside of East Africa.

Practically, we take advantage of this unique data structure to define a split-off household as any household that contains at least one member from the original roster and an extended family network as a network of split-off households, all originating from the same baseline household. Each household interview probed, through a network roster, for relationships and interactions between the current household and any split-off household. For example, if the original members have split into three different households, the network consists of these three households and each household is asked questions about the other two, giving us a data set of 6 dyadic observations.

What makes these data so particularly suited to shed light on our research question is, first, that we have data on both sides of each pair of linked households in the networks. That allows us to cross-check the beliefs held by one household about the other with reality as recorded in the questionnaire. Secondly, this survey is one of the few that tracks respondents outside the immediate vicinity of the enumeration area. While there are undoubtedly information asymmetries between households residing in the same locality, these may be more subtle and harder to measure. In our geographically disperse networks information asymmetries are salient - and below we will see that physical distance is strongly correlated with the degree of misperception between two households.

In empirical applications networks are typically self-defined, with questionnaires probing each respondent for a list of network partners. Our network definition is quite different as it is based on membership in a household 18 years ago. Our definition has the distinct advantage of being well-defined and exogenous, alleviating econometric concerns related to sampled networks (**?**). Attrition aside, we have complete networks defined in this way. Of course, as is nearly always the case in the literature, this network includes only a subset of households to or from whom transfers are sent or received. [3] We can quantify the share of the transfers that we are capturing by comparing the within-network transfers to transfers coming from outside the network. We see that for households receiving transfers, 51% of the donors are from within the extended family network as defined in this paper. Similarly, 51% of the total value of the transfers received is from within that network.

[2]Given the very long period of 16 years between the surveys the attrition rate in KHDS 2010 is extremely low relatively to other panels (Alderman et al. 2001)).

[3]In fact, we do not know of any matched survey with data on the comprehensive transfer network.

After dropping households that did not split or have missing or incomplete interviews, our sample consists of 3,173 households, from 712 families, yielding 13,808 unique within-family pairwise combinations of households. The principal strategy for tracing people from the original KHDS household rosters was to obtain their contact details through interviewed relatives. We should expect attrition rates to be higher among households that have infrequent contact with their family members. Indeed, out of all dyads in our sample 53% communicated at least once in the month preceding the survey, while for 5% the last communication was over 5 years ago. By contrast, the reports from interviewed households about untraced households (which constitute dyads that are dropped from the analysis) show that only 26% of such pairs communicated within the last month and 23% over 5 years ago. In what follows, then, it is useful to keep in mind that we are likely looking at a somewhat more connected set of family members who are not living in the same location.

Some (but not all) of the asymmetric information questions were skipped for split-off households residing in the same location as the respondent. Wherever the analysis below makes use of these skipped questions, we revert to a subsample of 9,032 dyads, all living in different locations, and encompassing 2,807 households within 613 extended family networks. Within this subsample 41% of dyads communicated within the past month and 7% within the past year.

Finally, significant resources were spent on collecting detailed consumption data on each interviewed household. The questionnaire included extensive food and non-food consumption modules, carefully designed to maintain comparability across survey rounds. For seasonal consumption items the recall period was 12 months to ensure comparability. The CAPI application automatically linked the consumption section to the agricultural section so the interviewer could probe carefully for consumption from home-produced foods. The final consumption aggregate includes purchased and home-produced food, as well as food eaten outside of the household. It contains 51 food items and 27 non-food items. The aggregates are temporally and spatially deflated using data from a price questionnaire included in the survey. Consumption is expressed in annual per capita terms using 2010 Tanzanian shillings.[4]

[4] A full description of the consumption aggregate is available at http://www.edi-africa.com/research/khds/introduction.htm

Table 1: Asymmetric information.

	Underestimate	**Spot-on**	**Overestimate**	**DK**
Completed O'level	0.029	0.920	0.022	0.029
Has formal job	0.054	0.872	0.008	0.066
Owns house	0.088	0.748	0.095	0.068
Owns land	0.120	0.761	0.048	0.071
Owns livestock	0.059	0.745	0.072	0.123
Owns phone	0.096	0.725	0.075	0.105
Owns TV	0.080	0.720	0.044	0.156
Owns motorized vehic	0.040	0.667	0.142	0.151

Notes: Comparing actual realizations to beliefs held by extended family members. Cells indicate the proportion of observations. Completed O'level means having completed the first four (out of six) years of secondary education. N=9,032.

3 Quantifying Asymmetry of Information

3.1 Beliefs about assets, education and employment

We can measure the extent to which extended family members are (mis)informed about each other by cross-checking the beliefs of any household i about educational attainment, employment and asset ownership of household j with the information in household j's questionnaire. We can do this over 8 different items listed in Table 1.[5] Household i can underestimate, correctly estimate or overestimate the status of household j; household i could also answer that it does not know the educational attainment, employment status or assets owned by household j. Table 1 gives frequencies of these cross-reports and already reveals some interesting patterns. Most underestimates of assets occur with respect to land and phones, while most overestimates occur with respect to vehicles. Educational attainment and employment have the most correct perceptions. Note, however, how very few people overestimate the employment position of their relatives, while relatively more underestimate (i.e. think their family members don't have a formal job, while in fact they do).

Ultimately, we are interested in measuring what i's perception on these 8 items tells us about

[5] Perceptions of educational attainment and occupation were collected at the individual level, for each original panel member (people who were member of the baseline household) currently member of j. The perceptions on the 6 asset were asked the j-household level.

i's perception of j's wealth, and to what extent and in which direction i misperceives j's true wealth. A first, rough such summary measure of asymmetry of information could be the simple sum of perceptions on the 8 items above adding up overestimates (set to +1), underestimates (set to -1), correct responses (set to 0) and don't know responses (set to 0). The first panel in Figure 1 shows the distribution of this sum, overlaid with a normal density curve, scaled to have its mean (-0.57) and standard deviation (1.32). This measure would lead us to conclude that, while there's a concentration (44%) of responses at zero, there's also a slight tendency to underestimate. Some of the lumping at zero happens because we count don't know responses as zeros. However, even on the subsample of observations that do not have any don't know responses, we still have 41% of dyads with perfect knowledge on the eight characteristics in Table 1.

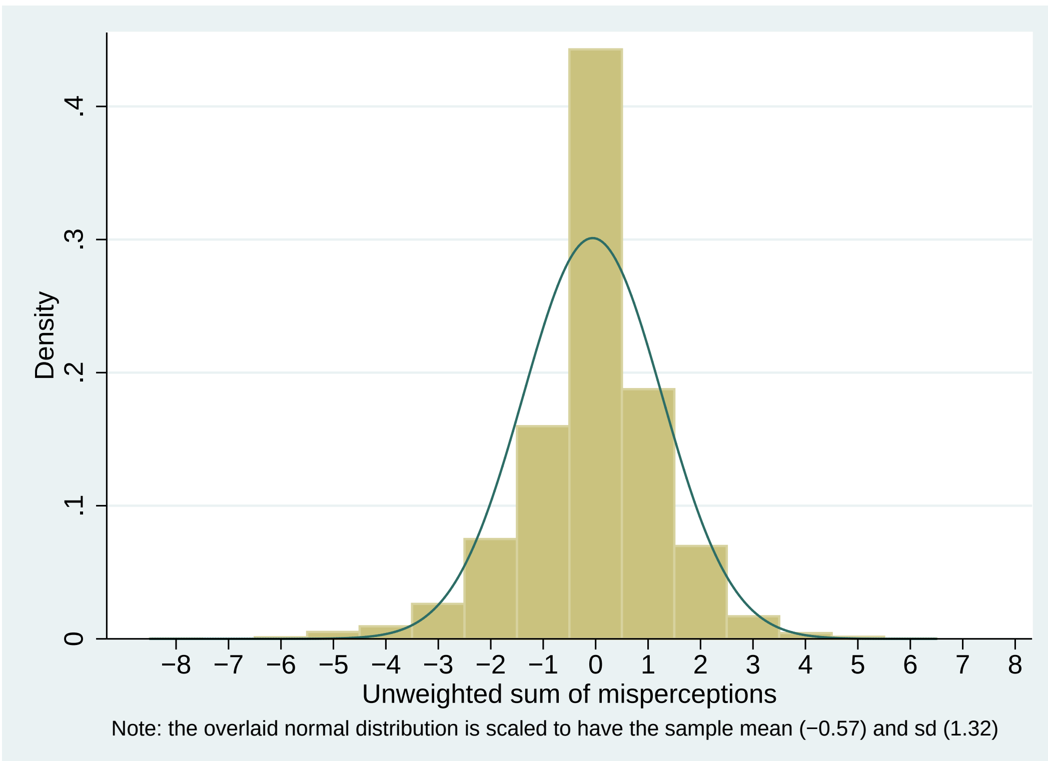

Figure 1: Simple sum of misperceptions (N=9,032)

There are a number of problems with the above simple sum approach. First, it is assumed that all items carry the same weight in the measure. Second, it is assumed that the weight of each item is separable from the household's characteristics, and from other items. Third, it is assumed that all items signal something positive about the wealth of the individual. An example of a violation of these last two assumptions would be if livestock ownership signals high income for rural households, but low income for urban households. In the next section we alleviate these concerns with a novel measure of asymmetric information: a weighted sum of misperceptions and their interactions, with weights set according to their correlation with household income.

3.2 Perceived Consumption and Misperceptions

Let $\mathbf{A}_j$ denote the profile of actual ownership of assets, education and occupation for household j, including any relevant interaction effects among assets and with urban or rural location of the household. This is the profile over which i expresses beliefs $\mathbf{B}_{ij}$. In addition, let X_j denote a vector of basic characteristics of household j that we would think are public knowledge among relatives.

Given i's belief about the profile of ownership of j, what can we infer about i's estimate of j's consumption? The view that we take here is that by observing households around him, i has learned the joint distribution of ownership profiles A_j and household welfare (log consumption per capita) $ln(c_j)$ conditional on household characteristics X_j and A_j.

Hence, we first estimate a consumption regression among our households

$$ln(c_j) = \mathbf{A}_j\alpha + \mathbf{X}_j\beta + \epsilon_j \tag{1}$$

where c_j is the actual per capita consumption for household j. To clarify the exposition we do not write all the interaction terms between X_j and A_j, which, if controlled for in the regression 1, can easily be carried forward into equations 3 and 4.

Retrieving the coefficients estimated $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ from (1), we can then use the characteristics $\mathbf{X}_j$ and i's beliefs about j's assets $\mathbf{B}_{ij}$ to construct measures of i's perception of j's consumption, $ln(C_{ij})$. Let

$$ln(C_{ij}) = \mathbf{B}_{ij}\alpha + \mathbf{X}_j\beta. \tag{2}$$

One benchmark is to think that households are much better informed than we are about all unobservable characteristics, including temporary shocks, that affect the income of their relatives. Hence, at the one end of the spectrum, we can assume that j's relatives are perfectly informed of

ϵ_j, in which case i's perceived consumption for j is

$$P_{ij} = \ln(C_{ij}) + \epsilon_j.$$

In this case, if household i holds perfect knowledge on household j's assets then the predicted consumption equals the actual consumption.

Another benchmark is to assume that household i uses only $\mathbf{X}_j$ and her beliefs about j's assets $\mathbf{B}_{ij}$ in forming her estimate of household j's per capita consumption $\mathbf{C}_{ij}$. This assumes that household i has no additional information about household j, over and above $\mathbf{X}_j$. In this case, i's perceived consumption for j is

$$P'_{ij} = \ln(C_{ij}).$$

Most likely, the truth lies somewhere in between these two estimates.

Using these two benchmarks, we can create the following two measures of misperceptions – the difference between i's perceived income for j and j's actual income:

$$\Omega_{ij} = P_{ij} - ln(c_j) = (\mathbf{B}_{ij} - \mathbf{A}_j)\,\alpha. \tag{3}$$

and

$$\Omega'_{ij} = P'_{ij} - ln(c_j) = (\mathbf{B}_{ij} - \mathbf{A}_j)\,\alpha - \epsilon_j. \tag{4}$$

Our measure of misperception Ω_{ij} is a weighted sum of the difference in believed and actual occupation, education and assets. As such it is well suited to measure misperceptions of medium-term income, rather than asymmetry of information regarding temporary shocks. In contrast, Ω'_{ij} might measure not only misperceptions on medium term income but be also affected by beliefs regarding temporary shocks, which are part of ϵ_j. For example, a temporary positive consumption shock to j, $\epsilon_j > 0$, with constant $\mathbf{A}_j$ will lower Ω'_{ij} – making it more likely for us to conclude that i underestimates j's wealth – but will not affect our measurement of Ω_{ij}.

The log specification conveniently implies that Ω_{ij} is a good approximation of the percentage by which i overestimates ($\Omega_{ij} > 0$) or underestimates ($\Omega_{ij} < 0$) j's consumption.

3.3 Measuring Weights for Perceived Consumption

We populate $\mathbf{X}_j$ with variables describing the gender and age of the household head, 8 variables capturing the age-sex composition of the household and a dummy indicating whether the household lives in a rural or urban area. These variables, described in Appendix, are assumed common knowledge.

When predicting household i's beliefs about household j we need to decide how to treat don't know (DK) answers to the assets, educational attainment and employment questions. In these regressions we take the conservative approach of replacing DKs with location-specific sample means, depending on whether household j lives in an urban or rural area.

We use a recursive method to establish which variables enter $\mathbf{B}_{ij}$. We start with a regression, shown in column 1 of Table 12 in Appendix 1, that includes all variables from Table 1: educational attainment and employment of the panel member (i.e. who belonged to the origin household, which was interviewed in the 1991-1994 survey), and household ownership of a house, land, livestock, phone, TV or motorized vehicle. The initial regression includes a set of interaction terms between all of these variables and the urban-rural dummy, to capture location-specific correlations between assets and wealth. We also interact the two largest and immobile assets, house and land, with other assets to explore complementarities. The second regression, shown in the second column of Table 12, retains only those elements of $\mathbf{B}_{ij}$ with $t > 1.5$ and the third regression in column 3 iterates the same procedure. After 3 iterations all coefficients of $\mathbf{B}_{ij}$ have $t > 1.5$ – a desirable feature of our final regression, as we want to avoid insignificant variables influencing the calculation of Ω_{ij}. The final weights obtained through the third regression are reported in Table 2, which also shows the high predictive power of the regression, explaining 58 percent of the variation in consumption.

The final specification highlights the importance of allowing for interaction effects when establishing the weights. We see, for example, that phone ownership shows less correlation with wealth in urban areas, while livestock ownership has no correlation with wealth in urban areas. Vehicle ownership, by contrast, is a stronger predictor of high consumption in urban areas. There is also a positive and significant interaction effect between house and vehicle ownership. Perhaps surprisingly, land ownership is dropped in the final regression as it is insignificant in levels and all interactions.

Applying the weights from Table 2 to Equation (3) and (4), we can calculate Ω and Ω', for each $i - j$ pair. Figure 2 shows kernel density estimations for both measures. The mean (standard

Table 2: Weights in Ω

	Coefficient
HH located in urban area	0.243***
	(6.313)
Panel member finished O level	0.176***
	(5.874)
Panel member has a formal job	0.177***
	(5.460)
Owns house	-0.083***
	(-3.686)
Owns livestock	0.146***
	(4.924)
Owns phone	0.308***
	(14.191)
Owns TV, video equipment or camera	0.313***
	(11.987)
Owns motorbike, car, truck or other vehicle	0.147*
	(1.756)
Owns house * owns vehicle	0.237***
	(3.026)
Urban * owns livestock	-0.136**
	(-2.129)
Urban * owns phone	-0.115***
	(-2.579)
Urban * owns vehicle	0.141**
	(2.284)
Adjusted R-squared	0.576
N	3173

Notes: Final weights in Ω determined through recursive estimation of Equation (3). All iterations are given in Table 12 in Appendix 1. t statistics between brackets under the coefficient. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

deviation) of 0.008 (0.309) for Ω and 0.004 (0.527) for Ω': on average there is a very slight, 0.3% to 0.8% overestimation of other's wealth, but there is no indication of any major systematic overestimation or underestimation of wealth. The average of the absolute values of Ω and Ω' are 0.21 and 0.42, meaning that people are, on average, 20 and 42 per cent mistaken, respectively. The kernel smoother from Figure 2 obfuscates the fact that Ω equals exactly zero in 27% of dyads, where all guesses were correct. This is a much lower percentage than what we observed in a simple sum, where any over and under estimations can cancel each other out exactly. The Ω' distribution does not exhibit this lumping at zero because it subtracts the ϵ_j term in its calculation. For the same reason the distribution of Ω' has a much larger spread than that of Ω.

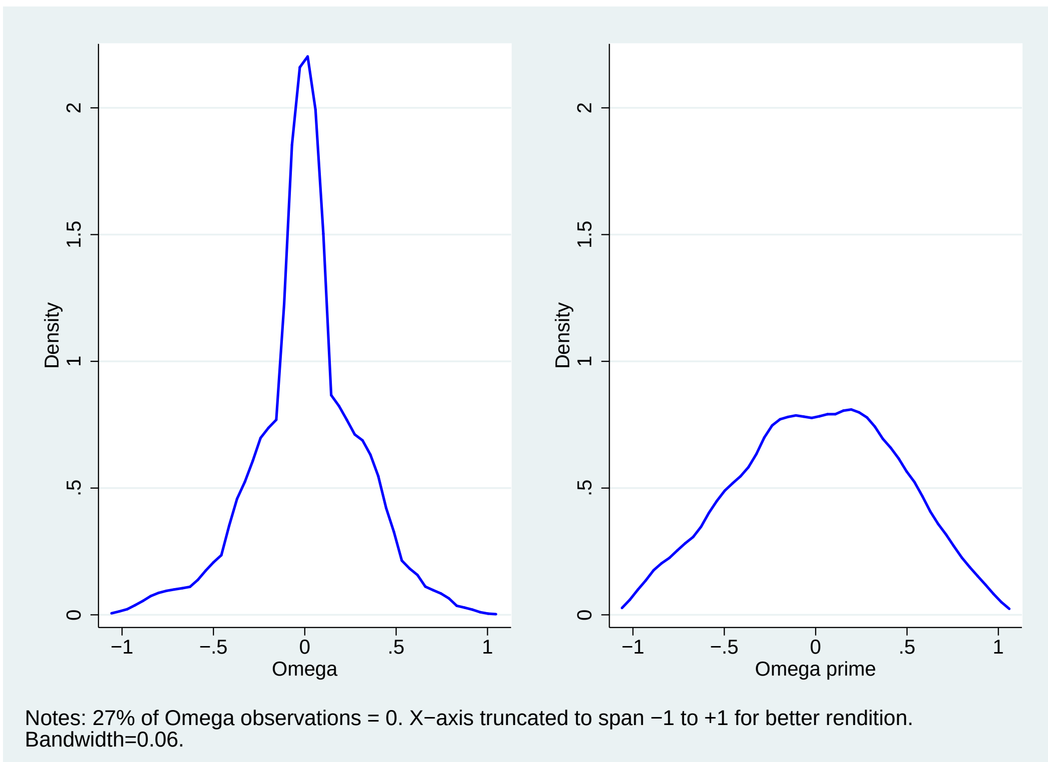

Figure 2: Distribution of Ω and Ω' (N=9,032)

3.4 Validation

We validate Ω and Ω' by comparing them with a completely different measure of asymmetric information. For each of its network partners, any household i was asked to imagine a nine-step ladder where the top of the ladder, step 9, represents the best possible life and the bottom, step 1, represents the worst possible life. Household i was then asked to rank each household j in his network on the ladder. In Table 3, we show that, using household i fixed effects, the higher household i places household j on the ladder, controlling for household j's actual consumption, the higher Ω_{ij} and Ω'_{ij} are. In other words, for 2 relatives with the same actual consumption household i's misperception of their wealth is highly correlated with where he differentially places each on the ladder. The strong correlations between our misperception measures and these subjective perceptions give confidence that Ω and Ω' are indeed capturing latent beliefs and are not merely noise. In Section 4 we will further validate our misperception measures by verifying how they co-vary with variables that measure the fluidity of information flows between the two nodes. Finally, when testing the model in Section we will use the ladder as an alternative measure of misperceptions to verify the robustness of the results.

Table 3: Comparing Ω to subjective perceptions.

	Ω_{ij}	Ω'_{ij}
i places j on bottom 3 rungs of 9-step ladder	-0.068***	-0.110***
	(-7.191)	(-9.156)
i places j on top 3 rungs of 9-step ladder	0.102***	0.231***
	(5.412)	(9.596)
i gives DK answer to ladder question regarding j	-0.056***	-0.038**
	(-4.226)	(-2.252)
log of j's actual consumption per capita	-0.121***	-0.562***
	(-21.872)	(-79.568)
Constant	1.637***	7.463***
	(22.265)	(79.532)
N	9032	9032

Notes: Household i fixed effect regression of Ω and Ω' on 3 ladder dummies indicating where household i places household j on a 9-step ladder.

4 Description of Misperceptions

4.1 Proximity and Misperception

It is very natural to expect individuals who are closer to each other, physically or socially, to have better information about each other. Table 4 relates misperceptions to a number of proximity variables through probit regressions. In the first column, the dependent variable is whether Ω_{ij} is zero, which requires that all the beliefs of i regarding j are correct and occurs in 27% of all dyads. In the second column the dummy is whether Ω'_{ij}, which is never exactly zero, falls in the middle 27% of the distribution. That way both regressions have an equi-proportional number of 0 and 1 observations.

In line with our expectations we see that the accuracy of perceptions, and in particular as measured by Ω_{ij}, increases with proximity variables such as physical distance (migration status and the kilometers of geographic distance), genetic distance (whether a parent-child link exists across the two households) or social distance (whether they communicated in past 2 years and whether they recently shared a meal together). We also see that there is less accurate information about extended family members living in urban areas, controlling for migration status.

The results also give further credence to our measures of information asymmetry by showing that they correlate as expected with some key variables and are not just capturing noise. If seen as a validation exercise, Ω clearly outperforms Ω'. Therefore, in what follows we will report results for Ω only. All our main results remain robust to using Ω' instead of Ω.

4.2 Perceptions within and across groups

Next, we ask whether there is any evidence of systematic perception errors within or across specific groups. We start by defining the poor as those falling in the lowest per capita consumption quartile and the rich as those falling in the highest quartile. The full line in the first panel of Figure 3 shows the perceptions of the poor about extended family members that are also poor. This can be compared to the dashed line which are the perceptions of the poor about their rich extended family members. The second panel takes the perspective of the rich and shows their perception of the poor versus the other rich. These within and across group perceptions show that the best information is by the poor about the other poor. The rich have much worse information about the

Table 4: Correct expectations and distance.

	Ω_{ij}	Ω'_{ij} **in middle 27% of distribution**
Distance between HHs (km)	-0.0002***	-0.0000
	(-10.100)	(-0.356)
Moved away from village and vicinity	-0.0656***	0.0179*
	(-6.557)	(1.786)
j located in urban area	-0.0502***	-0.0182*
	(-4.770)	(-1.705)
Parent-child link	0.0351***	0.0181
	(2.743)	(1.399)
i and j communicated in the past 2 years	0.1462***	-0.0138
	(12.953)	(-1.000)
Number of years since i and j last lived together	-0.0009	-0.0002
	(-0.978)	(-0.216)
Shared at least one meal in the past month	0.0506***	0.0276*
	(3.265)	(1.749)
N	9032	9032
Percent of observations with LHS = 1	27%	27%

Notes: Probit regressions of correct expectations. Marginal effect reported. t statistics in brackets under the coefficient. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Regressions include an indicator variable for 36 missing km distance observations.

poor compared to the poor themselves (judging by the lower peak at zero), but still slightly better on the poor than on other rich people. The omega distributions about the rich lie to the left of those about the poor, irrespective of whose perspective we take. This could simply be a reflection of uncertainty about the wealth levels, making it more likely to underestimate a rich person and overestimate a poor person.

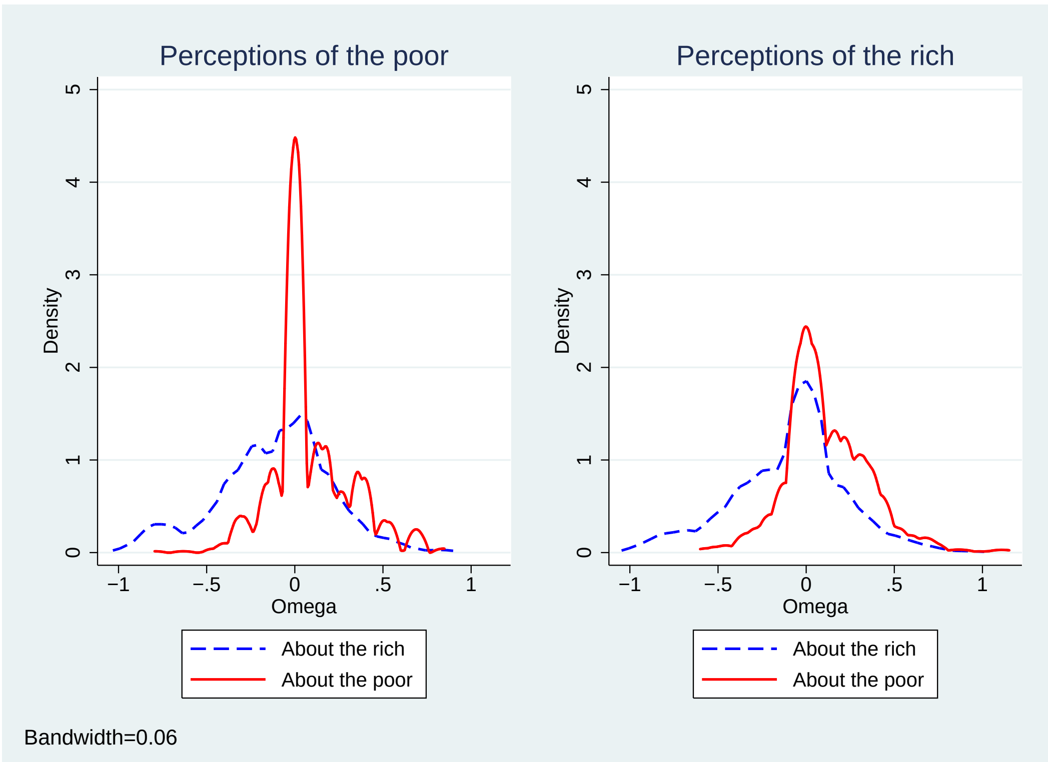

Figure 3: Kernel density functions of Ω, perceptions by wealth status (N=9,032).

Next, we split the population up by two characteristics that we can plausibly assume to be known within extended family networks: whether a household migrated or not and whether they live in an urban or a rural area. In Figure 4 we look at the perceptions of non-migrants (people still living in the baseline village or a neighboring village) about migrants versus other non-migrants (first panel) and perceptions of migrants about family members who remained at home versus other migrants in the network. Figure 5 does the same for urban versus rural dwellers.[6]

[6]It is useful to remember here that, as noted in the data section, the questionnaire skipped the cross-reports when people were living in the same location. That means, for example, that the non-migrants reports about each other

Systematic under or over estimation across whole groups, delineated by a known characteristic, is less readily explained. Finding systematic shifts to the left or right in the mass of the distribution of Ω in Figure 4 or 5 could be evidence of mass deception by migrants or urban dwellers about their actual asset holdings. Indeed, a number of recent studies suggest that migrants might want to underreport their income to avoid sending remittances (see McKenzie, Gibson, and Stillman (2012), Jakiela and Ozier (2012), Ambler (2013) and Seshan (2013)). The question is whether this deception could be sustained. One might expect individuals to anticipate such dissimulation and to have correct expectations on average.

Figure 4 shows that there is no major systematic underestimate nor overestimate of migrants' wealth: Ω is centered around 0 in the sense that the median of all these distributions is 0. Still, these distributions have different shapes on either sides of the 0-center with non-migrants underestimating migrants by about 1.8% on average and migrants overestimating the wealth of non-migrants by about 5.2% on average. These differences are small, but consistent with McKenzie et al. McKenzie, Gibson, and Stillman (2012) who find that prospective migrants from Tonga under-estimate their potential earnings in New-Zealand, and especially so if they have relatives who migrated to New Zealand (suggesting that they probably underestimate the earnings of their relatives), and with Serror Serror (2012) who finds that Senegalese migrants have substantial asymmetry of information regarding the asset holding of their family members.

A more pronounced difference is that non-migrants have the best information about each other, as evidenced by the high peak in the solid line of panel A in figure 4. Also migrants have better information about non-migrants than they do about other migrants, but the difference is smaller. Figure 5 shows that rural dwellers have much better information on other rural dwellers, compared to urban dwellers. A thinner spread could reflect better knowledge about circumstances and/or reflect a lower actual spread of asset-wealth. For example, there could be more accurate knowledge about the life of non-migrants as they have remained living in the place where migrants originate from too. Or non-migrants actual asset-wealth may display a lot less variation.

5 Model of Misperceptions and Transfers

In this Section, we present simple static models relating income, misperceptions of income and transfers. In a static model, there are three possible motivations for transfers: altruism, exchange

are not from people living in the same village, but rather from network members in neighboring villages

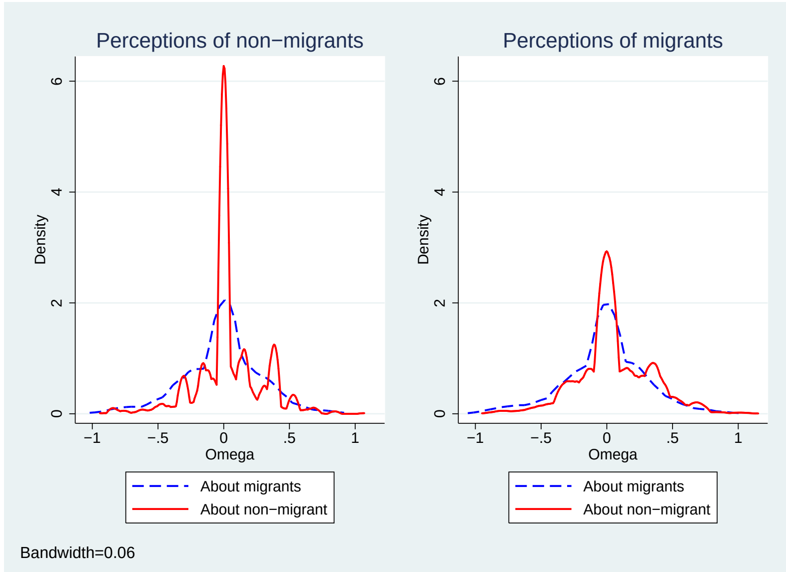

Figure 4: Kernel density functions of Ω, perceptions by migration status (N=9,032).

and pressure. These are incorporated in the model below. We show these motives have different predictions regarding the correlation between income, misperceptions and transfers.

The static nature of the model follows from the static nature of our data and the fact that misperceptions are asset based. To be sure, asymmetry of information regarding short term shocks and the resulting dynamic effects in transfers (see Cole and Kocherlakota (1999), Attanasio and Pavoni (2011) and Hauser and Hopenhayn (2008) among others) may also be present, but would be difficult to study in this context.

Preferences and Income:

Consider two individuals, a recipient R and a donor D. A first possible motive for transfers is altruism: the donor potentially cares not only about her own utility of consumption, but also about the recipient's. Denote by $u_i(c)$ the utility of consumption of $i \in \{R, D\}$, with the usual properties that $u' > 0$ and $u'' < 0$, and denote by $\alpha_D \in [0, 1)$ D's altruism, that is the weight that the donor puts on the recipient's utility of consumption (following Becker (1974)).[7].

[7] See also Stark (1995).

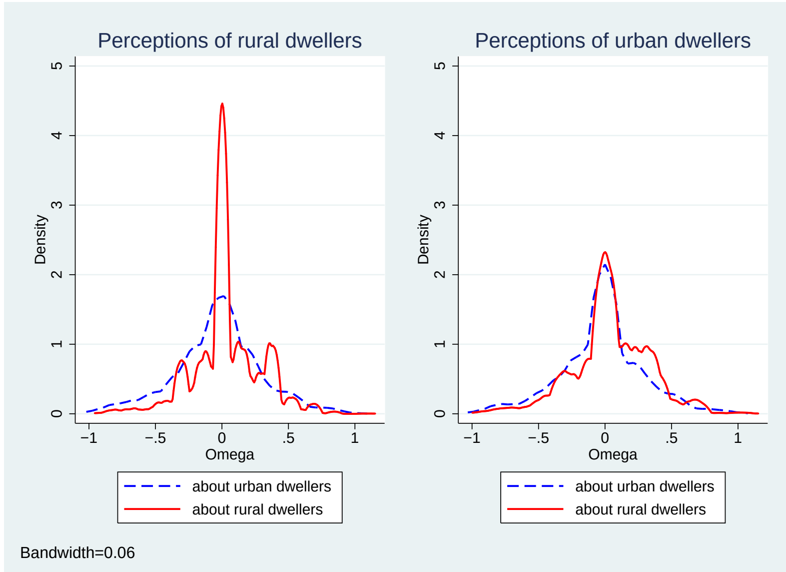

Figure 5: Kernel density functions of Ω, perceptions by location (N=9,032).

The recipient R and donor D's utilities are then $v_R = u_R(c_R)$

$$v_D = u_D(c_D) + \alpha_D \mathbb{E}_D u_R(\widetilde{c}_R) \tag{5}$$

where $\widetilde{c}_R$ is R's consumption as perceived by D. Note that we could easily assume that the recipient is also altruistic towards the donor.

Income and Information:

We assume that R and D's actual incomes are private information, though the following income distributions are common knowledge. The donor D's income y is either low (L) with a probability $1 - q_D$ or high (H) with a probability q_D, $L < H$ and $q_D \in (0, 1)$. The recipient R's income x takes a low value (ℓ) with probability $1 - q_R$ or a high value (h) otherwise, $\ell < h$ and $q_R \in (0, 1)$.

Let $i \in \{R, D\}$ be one party and $j \neq i \in \{R, D\}$ be the other. Individual i's beliefs about j's income are based on j's actual probability of having a high income, but will also reflect any information about the actual realization of j's income that i receives. Assume that i receives a signal $s^j \in (0, 1)$ about j's income m drawn from the conditional distribution $f_j(s|m)$. The realization of the signals is common knowledge. We assume that the conditional distributions satisfy the monotonic likelihood property, so that high values of the signal are relatively more likely when income is higher, but also that at the extreme, these signals are almost perfectly informative:

[$S1$] $\frac{f_R(s|h)}{f_R(s|\ell)}$ and $\frac{f_D(s|H)}{f_D(s|L)}$ are strictly increasing in s,
[$S2$] $\lim_{s\uparrow 1} \frac{f_R(s|\ell)}{f_R(s|h)} = \lim_{s\uparrow 1} \frac{f_D(s|L)}{f_D(s|H)} = 0$ and $\lim_{s\downarrow 0} \frac{f_R(s|h)}{f_R(s|\ell)} = \lim_{s\downarrow 0} \frac{f_D(s|H)}{f_D(s|L)} = 0$.

After using the signals and Bayes rule to update their beliefs, the posterior beliefs that the recipient and the donor hold about each other are given by

$$\pi^R_{s_R} = \frac{q_R f_R(s_R|h)}{q_R f_R(s_R|h) + (1-q_R) f_R(s_R|\ell)} \; and$$
$$\pi^D_{s_D} = \frac{q_D f_D(s_D|H)}{q_D f_D(s_D|H) + (1-q_D) f_D(s_D|L)},$$

where π^j_s is the probability that i assigns to j having a high income after observing signal s.

Pressure: Although more rarely studied, pressure seems to be an important determinant of transfers. There is a large literature describing the pressure under which many households in developing countries find themselves to assist relatives (see Chort, Gubert, and Senne (2012) and Platteau (2012) among others). Individuals may be able to shame relatives who fail in helping them in times of known need, or make them feel guilty. Alternatively recipients could use their community to exert this pressure and retaliate against relatives who fail to transfer enough. Pressure can also be through a loss of social status if one fails to transfer enough. This pressure might be available to recipients only at certain times, for instance when the recipient has a well known need (she suffers an observable shock, the school fees are due, etc).

To model this pressure motive, we assume that R can commit on imposing a utility cost $p \in [0, P]$ onto D.

Exchange: Finally, another possible motive for transfers, in particular for migrants, is quid-pro-quo. As discussed by Cox and Fafchamps (2008) and Rapoport and Docquier (2006), private transfers might be given in exchange for goods or services provided by the recipient. This could be help with young children, old-age support or maintaining property rights for migrants. Assume that, at times, the recipient is in position to provide a service of utility value v to the donor at a utility cost c.

Altruism

Consider first a situation in which altruism is driving the transfers: $\alpha_D > 0$, $P = 0$ and $v = 0$. The recipient R has no credible way to signal her income. Hence, for a given realization of her income y and the signal $s^R = s$, D chooses to make a transfer t to R that maximizes

$$u_D(y-t) + \alpha_D[\pi_s^R u_R(h+t) + (1-\pi_s^R)u_R(\ell+t)]. \tag{6}$$

D's choice of transfer t^* clearly depends on his own income as well as on his posterior beliefs regarding R's income: a positive transfer t^* is strictly increasing in y and decreasing in π_s^R. Hence, the altruism model predicts a positive correlation between transfers and the donor's actual income, and a negative correlation between transfers and the donor's perception of the recipient's income. The latter implies a negative correlation between transfers and the donor's misperception of the recipient income controlling for the actual income and a negative correlation between transfers and the recipient's income controlling for the donor's misperception.

Since there is a one-to-one correspondence between the donor's income and the transfer that she chooses when a positive transfer is made, upon receiving a transfer from D, R would know D's realized income. Hence, no correlation is predicted between the transfers and the recipient's misperception of the donor's income.

Note that the same predictions would apply if the recipient was also altruistic towards the donor although an interesting signaling game can arise in this case as shown in Genicot (2014).

Pressure

Now, let's study the case without altruism $\alpha_D = 0$ and without services $v = 0$, but in which transfers are driven by the possibility of pressure: $P > 0$. Since the cost for D of making a given transfer is decreasing in her income, R can make use of pressure not only to receive a transfer but also to get D to reveal her real income.

Indeed, given his income x, R offers a menu to D of transfers t and contingent pressure $p(t)$:

− a transfer of T_H or more implies no pressure;

− a transfer of T_L implies pressure $\underline{p}$; and

− any other transfer is associated with pressure $\overline{p}$.

The offer is designed so that the donor chooses to give T_H and face no pressure when her income is high, while she chooses to give T_L and face pressure $\underline{p}$ when her income is low.

Incentive constraints: To simplify notation, denote by $V_y(T)$ D's utility if her income is y and she makes a transfer T:

$$V_y(T) = u_D(y - T). \tag{7}$$

The incentive constraints for both type H and type L to comply with the offered menu are

$$V_H(T_H) \geq \max_t \{V_H(t) - p(t)\} \tag{8}$$

$$V_L(T_L) - \underline{p} \geq \max_t \{V_L(t) - p(t)\} \tag{9}$$

where $p(t)$ is the pressure triggered by the scheme. Clearly, in the absence of altruism, D's preferred transfer to R would be 0 while R would like to receive as much as possible. Hence, R uses the highest pressure as a threat $\overline{p} = P$ and, if D has a low income, her preferred deviation would be to make no transfer.

It follows that the incentive constraint for type L in (9) becomes,

$$V_L(T_L) - \underline{p} \geq V_L(0) - P. \tag{10}$$

Now, for type H, the relevant constraint ensures that H does not want to pretend to be L:[8]

$$V_H(T_H) \geq V_H(T_L) - \underline{p}. \tag{11}$$

Given his income x and the signal s received, the recipient R chooses $\underline{p}$, T_H and T_L to maximize

$$\pi_s^D u_R(x + T_H) + (1 - \pi_s^D) u_R(x + T_L) \tag{12}$$

subject to (10) and (11).

The following two types of offer (or contracts) are possible:

Pooling: R asks D to transfer an amount t^p and no pressure is effectively applied $\underline{p} = 0$, otherwise the maximal pressure $\overline{p} = P$ is applied. Hence, D makes the same transfer irrespective of her

[8]As usual, (10) and (11) imply that $V_H(T_H) - \underline{p} \geq V_H(0) - P$. This is shown in Appendix 2.

income $T_L = T_H = t^p$. Beliefs are therefore not updated and R never learns D's real income. The transfer t^p is such that a type L donor is indifferent between giving t^p and receiving the maximum amount of pressure, i.e. $u_D(L - t^p) = u_D(L) - P$.

or

Separating: R demands from D either a transfer of T_H in exchange for no pressure, or a lower transfer $T_L(< T_H)$ but with some pressure $\underline{p} > 0$. Any other transfer would result in maximal pressure. D chooses to transfer more when her income is high $T_H > T_L$ but he is subject to pressure $\underline{p} > 0$ when her income (and transfer) is lower. Since D's transfer varies with her income, R updates his beliefs and has full information ex-post. The exact values T_H, T_L and $\underline{p}$ depend on the probabilities π_s^D and on x.

Naturally, R is more likely to offer a separating contract when he receives a signal s^D that makes it more likely that D has a high income. This intuition is formalized in the following Proposition whose proof is in Appendix 2.

Proposition 1 *There is a cutoff value of the signal $\widetilde{s} \in (0,1)$ such that R offers a pooling contract if $s_D \leq \widetilde{s}$ and R offers a separating contract if $s_D > \widetilde{s}$ with T_H (T_L) increasing (decreasing) in s*

Proposition 1 tells us that if the donor's actual income is low, either
(i) $s_D > \widetilde{s}$ and R offers a separating contract, in which case D gives a low transfer and is subject to pressure and R's beliefs are correct ex-post; or
(ii) $s^D \leq \widetilde{s}$ and R offers a pooling contract with transfer $t^p > t_L$, in which case R overestimates D's income ex-post.

Similarly, if the donor's actual income is high, either
(i) $s_D > \widetilde{s}$: R offers a separating contract, in which case D gives a high transfer and R's beliefs are correct ex-post; or
(ii) $s^D \leq \widetilde{s}$: R offers a pooling contract with transfer $t^p < t_H$, in which case R underestimates D's income ex-post.

Table 7 illustrates these findings. First, there is a positive correlation between D's actual income and her transfer to R. Second, for a given income level of D (H or L) there exists a positive correlation between the perception of D's income by R (π^D) and the transfer from D to R: the more R thinks D has, the higher the transfer from D to R is. Because these are ex-post beliefs, this correlation takes into account the feedback mechanism through which transfers influence beliefs.

Table 5: Summary of the Pressure Model

D's income	Signal	Contract to D	Transfers to R*	Ω_{RD}
L	$s_D > \widetilde{s}$	separating	low (T_L)	$= 0$ (correct)
	$s^D \leq \widetilde{s}$	pooling	medium (t^p)	> 0 (overestimate)
H	$s_D > \widetilde{s}$	separating	high (T_H)	$= 0$ (correct)
	$s_D \leq \widetilde{s}$	pooling	medium (t^p)	< 0 (underestimate)

$* \; T_L < t^p < T_H$

In terms of the recipient's income, whether a higher x affects the contract offered and, if so, whether it encourages separating or pooling depends on the utility function. For instance, if the recipient utility function exhibits diminishing risk aversion then the recipient may be more likely to offer a separating contract when his income is high ($\widetilde{s}$ decreases) and, for any given signal s_D above the threshold, R may ask a higher T_H and a lower T_L when his income is high. Since a higher spread between T_L and T_H would be associated with a higher mean transfer, this would imply a small positive correlation between the transfers and the recipient's income.

Note that this effect could translate into a correlation of the same sign between the transfers and the donor's beliefs about the recipient's income, controlling for the actual incomes, but this correlation would be negligible. This is not only because the slightly higher transfer would come from an increase in the spread of the transfers, but also because this correlation would come only from the values of the signal s^D that reveals x's income: the values of the signal that are below the threshold when $x = \ell$ and above the threshold when $x = h$. For signals above the threshold when $x = \ell$, the transfers requested inform the donor of the recipient's income and there is no remaining misperception; while for signals below the threshold when $x = h$, the transfers requested are uninformative.

Similarly a utility function with increasing risk aversion could imply a small negative correlation between the transfers and the recipient's income, and a negligible correlation of the same sign between the transfers and the donor's beliefs about the recipient's income, controlling for the actual incomes.

Exchange

To study the exchange motive, assume that there is no pressure ($P = 0$) nor altruism ($\alpha_D = 0$), but that the recipient can provide a service of utility value v to the donor at a utility cost c, $c < v$.

The price of that service, the transfer, clearly depends on their relative bargaining power.

Denote as $\underline{t}(x)$, the lowest transfer that the recipient would accept to provide the service, given his income x:

$$u_R(x + \underline{t}(x)) - u_R(x) = c, \tag{13}$$

and as $\bar{t}(y)$, the highest transfer that the donor would pay for the service, given her income y:

$$u_D(y) - u_D(y - \bar{t}(y)) = v. \tag{14}$$

We assume that the relative value of the service (v/c) is sufficient such that the exchange is socially optimal: $\bar{t}(L) > \underline{t}(h)$. We follow Cox and Fafchamps (2008) and Rapoport and Docquier (2006) and consider in turn the two extremes: the case where the donor has all the bargaining power and the case where the recipient has all the bargaining power.

Exchange-D: Donor has the bargaining power

Assume that D gets to make a take-it-or-leave-it offer to R. This recipient's reservation price $\underline{t}(x)$ is clearly increasing in his income x. Hence, D essentially chooses between a) offering $\underline{t}(h)$ for the service, an offer that R always accepts, or b) offering a lower transfer $\underline{t}(\ell)$ that R accepts only when his income is low. Other offers are dominated. The optimal choice depends on D's income y and her beliefs regarding R's income π_s^R. D chooses a) if

$$\begin{aligned} & u_D(y - \underline{t}(h)) + v \geq \pi_s^R u_D(y) + (1 - \pi_s^R)(u_D(y - \underline{t}(\ell)) + v) \Leftrightarrow \\ & v \geq [u_D(y) - u_D(y - \underline{t}(h))] + \frac{1 - \pi_s^R}{\pi_s^R}[u_D(y - \underline{t}(\ell)) - u_D(y - \underline{t}(h))], \end{aligned} \tag{15}$$

and chooses b) otherwise. Higher π_s^R makes this inequality more likely to hold. For low values of the signal s^R, π_s^R is close to 0 and inequality (15) cannot hold, while for high values of the signal s^R, π_s^R is close to 1 and (15) is necessarily satisfied.

Proposition 2 *There is a cutoff value of the signal $\bar{s} \in (0, 1)$ such that D offers $\underline{t}(\ell)$ if $s_R \leq \bar{s}$ and D offers $\underline{t}(h)$ if $s_R > \bar{s}$.*

D offers $\underline{t}(\ell)$ when she receives a signal that the recipient's income is likely to be low ($s^R \leq \overline{s}$), and she offers $\underline{t}(h)$ when the signal indicates that recipient's income is likely to be high $s^R > \overline{s}$. Controlling for D's perception of R's income, the actual transfer is negatively correlated with R's actual income. And controlling for the actual realization of R's income (x), the correlation between the transfer and D's perception of R's income (π^R) is ambiguous: it depends on the value of x, positive for low values of x and negative for high values of x.

Table 6: Summary of Exchange-D scenario [3]

R's income	Signal	Offer to R	Transfers to R*	Ω_{DR}
ℓ	$s^R > \overline{s}$	$\underline{t}(h)$	$\underline{t}(h)$	> 0 (overestimate)
	$s^R \leq \overline{s}$	$\underline{t}(\ell)$	$\underline{t}(\ell)$	$= 0$ (correct)
h	$s^R > \overline{s}$	$\underline{t}(h)$	$\underline{t}(h)$	< 0 (underestimate)
	$s^R \leq \overline{s}$	$\underline{t}(\ell)$	0	$= 0$ (correct)

What about the donor's income y? A higher income makes inequality (15) more likely to hold. Richer donors are more likely to offer the high price $\underline{t}(h)$ so that the threshold $\overline{s}$ is smaller for richer donor. This would imply a positive correlation between the transfers and both the donor's income and the recipient's beliefs about the donor's income. These correlations come only from the realization of the signal that reveals y's income: the values of the signal that are between the threshold $\overline{s}$ for $y = H$ and the threshold $\overline{s}$ for $y = L$ and so are likely to be small.

Exchange-R: Recipient has the bargaining power

Now, assume that the recipient gets to make a take-it-or-leave-it offer to the donor. It is easy to check that the donor's reservation price $\overline{t}(y)$ too is increasing in her income. Hence, R essentially chooses between two options: a) demanding $\overline{t}(L)$ for the service, an offer that D always accepts, or b) demanding a higher transfer $\overline{t}(H)$ that D rejects when her income is low but accepts when her income is high. Other demands would be dominated by one of these two options. R's chosen option depends on his income x and his beliefs about D's income π_s^D. R chooses a) if

$$u_R(x + \overline{t}(L)) - c \geq \pi_s^D(u_R(x + \overline{t}(H)) - c) + (1 - \pi_s^D)u_R(x) \Leftrightarrow$$

$$[u_R(x + \overline{t}(L)) - u_R(x)] - c \geq \frac{\pi_s^D}{1 - \pi_s^D}[u_R(x + \overline{t}(H)) - u_R(x + \overline{t}(L))], \tag{16}$$

and chooses b) otherwise. The higher π_s^D is the more likely R is to ask $\bar{t}(H)$. Again we can see that when π_s^D is close to 0, inequality (16) is satisfied while it fails for value of π_s^D close to 1.

Proposition 3 *There is a cutoff value of the signal $s^* \in (0,1)$ such that R asks $\bar{t}(L)$ if $s_D \leq s^*$ and R asks $\bar{t}(H)$ if $s_D > s^*$.*

R asks $\bar{t}(L)$ when he receives a signal that the donor's income is likely to be low, and he offers $\bar{t}(H)$ when the signal indicates that donor's income is likely to be high. Controlling for D's income (y), the transfer is positively correlated with R's perception of D's income (π_s^D) and, controlling for R's perception of D's income, the transfer is positively correlated with D's income.

Table 7: Summary of Exchange-R scenario [3]

D's income	Signal	Demand to D	Transfers to R^*	Ω_{RD}
L	$s^D > s^*$	$\bar{t}(H)$	0	$= 0$ (correct)
	$s^D \leq s^*$	$\bar{t}(L)$	$\bar{t}(L)$	> 0 (overestimate)
H	$s^D > s^*$	$\bar{t}(H)$	$\bar{t}(H)$	$= 0$ (correct)
	$s^D \leq s^*$	$\bar{t}(L)$	$\bar{t}(L)$	< 0 (underestimate)

Again, there will be a correlation between the recipient's income x and the transfers only if it affects the scenario that he chooses. As a higher income for the recipient could make him more or less likely to select a) depending on the values of $\bar{t}(L)$ and $\bar{t}(H)$ and his utility function, this correlation, if there is one, could go in any direction. The same holds for the beliefs of donor about the recipient's income.

Predictions

Table 8 summarizes the predictions of the altruism, pressure and exchange models for our empirical section.

Note that we recognize that in the data some misperceptions are due to measurement errors/noise. Hence, when the model predicts no misperception at all and therefore no correlation could be calculated, we enter 0. If the misperceptions are just due to noise then one would expect a zero correlation.

Table 8 presents the predicted *partial* correlation between the transfers and the donor's and recipient's income and misperception. That is, it shows the predicted correlations between the transfers and the donor's (recipient's) misperception controlling for the income realizations and the other's misperception, and the predicted partial correlation between the transfers and the donor's (recipient's) income controlling for the misperceptions and the other's income. These correspond to the predictions regarding the regression coefficients in the next section.

To derive the partial correlation between the transfers and the donor's misperception in this table, we assume that that, in the pressure and exchange-R models, R chooses the same scenario for different values of her income. As explained earlier, if the type of contract chosen by R changes with his income, then this could create a small correlation between the transfers to R and Ω_{DR} whose sign depends on the utility function of the recipient, but is most likely insignificant.

Table 8: Predictions

Partial correlation between transfer and: Model	Ω_{RD}	Ω_{DR}	D's income	R's income
Altruism	0	$-$	$+$	$-$
Pressure	$+$	0	$+$	A
Exchange-D $\underline{t}(l)$ or $\underline{t}(h)$	$+(small)$	A	$+(small)$	$-$
Exchange-R $\bar{t}(L)$ /$\bar{t}(H)$	$+$	0	$+$	0

Notes: A = ambiguous.

Note that having access to a (partial) hiding technology should not affect the predictions, though it would attenuate some of these effects.

6 Asymmetric information and transfers

6.1 Main results

The survey collected data on amounts remitted, both in kind and cash. Over two thirds of households in our sample report remitting in cash or kind in the year preceding the survey. The average amount remitted, among those who did, was USD 35, representing, on average, 7% of consumption per capita and 2% of total household consumption for the remitting households. That average masks a wide distribution, with the top decile remitting USD 160, signifying 25% of their consumption per capita and 8% of their total household consumption.

Out of these transfers, 59% goes to recipients within the extended family network. For these within-network transfers we have full interview data on both the recipient and the donor, including the perceptions they have of each other and the degree to which these perceptions deviate from the truth. This allows us to test the predictions from Table 8 regarding the partial correlations between transfers (T_{RD}) received by any household R in the family network from any other family D in the same network and (i) the degree of misperceptions of the recipient about the donor (Ω_{RD} or Ω'_{RD}), (ii) the degree of misperceptions of the donor about the recipient (Ω_{DR} or Ω'_{DR}), (iii) the donor's income (I_D) and (iv) (iii) the recipient's income (I_R), with both income variables proxied by log consumption per capita.

We estimate these partial correlations through dyadic regressions of T_{RD}, whether or not D reported giving transfers to R in the year preceding the survey,[9] on the four correlates we are interested in. This brings with it a number of econometric challenges. The first is that we need to condition the correlations on the correct variables in order to get unbiased estimates. In the model we assume everything constant across both donor and recipient. We can implement this empirically by using a two-way fixed effect model, which includes a fixed effect for R and D, α_R and α_D, respectively.[10] Of course this does not capture any dyadic specific heterogeneity that may cloud these correlations. However, we also control for a set of observable dyadic characteristics describing the relationship between donor and recipient households to minimise such concerns.

The second econometric issue we address is to allow for correlations between transfers received by the same recipient or sent by the same donor, which would otherwise lead to biased estimates of the standard errors. To correct for this we use, in our preferred specification, the non-nested two-way clustering approach developed by Cameron, Gelbach, and Miller (2011) and implemented in Stata by Baum, Shaffer, and Stillman (2007).[11] Later on we test for robustness when allowing all transfers to be correlated within extended family networks and find very similar results.

Our preferred regression then is

$$T_{RD} = \beta_1 \Omega_{RD} + \beta_2 \Omega_{DR} + \mathbf{P}_{RD}\gamma_1 + \alpha_R + \alpha_D + \epsilon_{RD}, \qquad (17)$$

where ϵ_{RD} is an error term and $\mathbf{P}_{RD}$ is a vector of variables describing the relationship between the R-D household pair. This includes whether the heads of both households have the same religion, are from the same tribe, the geographic distance between the two households, whether a parent-

[9] 25% of dyads have $T_{RD} = 1$

[10] These models have been discussed by Mittag (2012), with De Weerdt (2004) providing an early application of two-way fixed effects in dyadic regressions for network analysis.

[11] An alternative method is provided by Fafchamps and Gubert (2007)

child relationship exists between the two households and the number of don't know responses to the asymmetric information questions each side of the dyad gave. Standard errors are two-way cluster-robust, with clustering on both R and D.

One of the variables of interest from the model predictions in Table 8 is the donor's income I_D, which is subsumed in the fixed effect α_D. In order to retrieve an estimate of the coefficient of I_D we drop α_D in Equation (17) and replace it with D's income and other characteristics. That is, we estimate:

$$T_{RD} = \beta_1 \Omega_{RD} + \beta_2 \Omega_{DR} + \beta_3 I_D + \mathbf{P}_{RD} \gamma_1 + \mathbf{Z}_D \gamma_2 + \alpha_R + \epsilon_{RD}, \tag{18}$$

where $\mathbf{Z}_D$ is a vector of household D characteristics, which includes the sex, age and years of education of the head of the donor's household as well as the number of household members that fall in each of eight exclusive and exhaustive age-sex categories (these together also control for household size).

Similarly, to retrieve an estimate on the I_R variable we drop α_R in Equation (17) and replace it with R's income and other characteristics:

$$T_{RD} = \beta_1 \Omega_{RD} + \beta_2 \Omega_{DR} + \beta_4 I_R + \mathbf{P}_{RD} \gamma_1 + \mathbf{Z}_R \gamma_2 + \alpha_D + \epsilon_{RD}. \tag{19}$$

It is worth recalling that we do not attach a causal interpretation to these coefficients. In fact the model explicitly allows for feedback mechanisms between the level of transfers and perceptions. For example, in a separating equilibrium, a high T_{RD} will cause beliefs Ω_{RD} to be revised upwards. We use the dyadic regression set-up as a convenient way to retrieve the partial correlations, measured by β_1, β_2, β_3 and β_4, in order to compare their signs to the predictions from Table 8.

We estimate Equations (17), (18) and (19) separately for both Ω and Ω', giving a total of six regressions, shown in Table 9. All equations are estimated using linear probability models.

The picture that emerges from this table is clearly that $\beta_1 > 0$, $\beta_2 = 0$, $\beta_3 > 0$ and $\beta_4 = 0$, which is consistent with either pressure or exchange-R in Table 8. The absence of a negative coefficient on the the recipient's income, β_4, allows us to reject both altruism and exchange-D as motives of transfers. We can also reject the hypothesis that β_2, the coefficient on Ω_{DR}, is smaller than zero, thus rejecting a second prediction of the altruism model.

The absence of negative correlation between transfers and the recipient's income is consistent with the finding of Lucas and Stark (1985) and Cox, Eser, and Jimenez (1998) who reject altruism

Table 9: Main Results, Partial correlations with transfers: Ω and income

	Equation					
	(17)	**(18)**	**(19)**	**(17)**	**(18)**	**(19)**
Ω_{RD}	0.093**	0.047**	0.056***			
	(2.24)	(2.49)	(2.74)			
Ω_{DR}	0.003	0.038	0.009			
	(0.08)	(1.60)	(0.58)			
Ω'_{RD}				0.093**	0.101***	0.056***
				(2.24)	(4.85)	(2.72)
Ω'_{DR}				0.003	0.037	0.007
				(0.08)	(1.59)	(0.38)
I_D		0.090***			0.167***	
		(7.54)			(8.22)	
I_R			0.007			0.011
			(0.73)			(0.66)
R^2	0.72	0.11	0.09	0.72	0.11	0.08
N	9032	9032	9032	9032	9032	9032

Notes: Fixed Effect Linear probability models with t values in parentheses under the coefficient. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

as a motive for remittances among migrants in Botswana and as a motive for transfers in Peru. In contrast, Kazianga (2006) finds some support for the altruistic motive among the middle income class in Burkina Faso, but not for low income levels, and in the US, Altonji, Hayashi, and Kotlikoff (1997) finds a negative correlation between in-vivo transfers and the recipient's income in support of the altruism remittance motive.

We can check that our findings are unaltered if we use different measures of the perceptions of income, coming from the ladder questions described in Section 3.4. We run slightly different versions of Equations (17), (18) and (19), where we include a measure for whether i places j on the lowest steps 1, 2 or 3 of the ladder, $\underline{L}_{ij}$, or the highest steps 7, 8, 9 of the ladder, $\overline{L}_{ij}$. Furthermore, the variables measuring the number of don't know responses to asymmetric information questions are replaced by two dummy variables indicating that don't know responses were given to the ladder questions by the donor or the recipient, respectively. The two way fixed-effects version then becomes:

$$T_{RD} = \delta_1 \underline{L}_{RD} + \delta_2 \overline{L}_{RD} + \delta_3 \underline{L}_{DR} + \delta_4 \overline{L}_{DR} + \mathbf{P}_{RD}\gamma_1 + \alpha_R + \alpha_D + \epsilon_{RD}, \tag{20}$$

with cluster-robust standard errors, with clustering on both R and D. Note that these FE specifications allow us to capture any fixed unobserved characteristic of the recipient or of the donor, which may systematically affect their relative perceptions of the positions of other households in their family network.

As above, to retrieve estimates of the coefficients on I_D and I_R, we estimate

$$T_{RD} = \delta_1 \underline{L}_{RD} + \delta_2 \overline{L}_{RD} + \delta_3 \underline{L}_{DR} + \delta_4 \overline{L}_{DR} + \mathbf{P}_{RD}\gamma_1 + \beta_3 I_D + \mathbf{Z}_D\gamma_2 + \alpha_R + \epsilon_{RD} \tag{21}$$

and

$$T_{RD} = \delta_1 \underline{L}_{RD} + \delta_2 \overline{L}_{RD} + \delta_3 \underline{L}_{DR} + \delta_4 \overline{L}_{DR} + \mathbf{P}_{RD}\gamma_1 + \beta_4 I_R + \mathbf{Z}_R\gamma_2 + \alpha_D + \epsilon_{RD} \tag{22}$$

respectively.

Table 10 shows the estimates of these three equations. To be consistent with the results in Section 6.1, we would expect to see that, controlling for D's actual consumption (through either α_D or I_D), lower perceptions of R about D's wealth are correlated with a lower likelihood of transfers from D to R, which implies $\delta_1 < 0$ and, symmetrically, that higher perceptions of R about D's wealth are correlated with a higher likelihood of transfers from D to R, that is $\delta_2 > 0$. Similarly as before, we expect , $\delta_3 = 0$, $\delta_4 = 0$, $\beta_3 > 0$ and $\beta_4 = 0$. Table 10 does indeed yield this pattern, except for the estimated coefficients δ_2 which turn out to be largely not significantly different from

Table 10: Partial correlations with transfers: ladder and income

	Equation		
	(20)	**(21)**	**(22)**
R places D low on ladder ($\underline{L}_{RD}$)	-0.026*	-0.051***	-0.015*
	(-1.72)	(-4.57)	(-1.71)
R places D high on ladder ($\overline{L}_{RD}$)	-0.032	0.025	-0.019
	(-1.01)	(1.01)	(-0.95)
D places R low on ladder ($\underline{L}_{DR}$)	0.001	0.017	0.005
	(0.04)	(1.60)	(0.51)
D places R high on ladder ($\overline{L}_{DR}$)	-0.041	-0.017	-0.032
	(-1.19)	(-0.73)	(-1.45)
I_D		0.078***	
		(7.75)	
I_R			0.004
			(0.50)
R^2	0.64	0.09	0.08

Notes: Fixed Effect Linear probability models with t values in parentheses under the coefficient. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. $N = 13,808$.

zero at conventional levels. This anomaly could be attributed to the fact that this category holds only 5% of the observations, but expanding the boundaries does not make a difference. All in all, we take these results as confirming the main results indicating that the power in the gift-giving relationship lies with the recipient of the transfer, consistent with either a pressure model or an exchange model in which the recipient holds the bargaining power.

6.2 Robustness checks using linear probability models

We now discuss four robustness checks of Equations (17), (18) and (19). Results are presented in Appendix 4.

First, the regressions presented so far use information on transfers remitted to a recipient, as reported by the donor. In our data all transfers between two households are reported by both parties. Table 15 takes advantage of this feature to check that results are robust to using the recipient's report instead of the donor's report.

Second, in Table 16 we check for robustness to changing the level of clustering. We replace the two-way clustering with clustering at the level of the extended family network.

Third, one may worry about spurious correlations between transfers and incomes since transfers out are not consumed by the donor (diminishing her income, as proxied by total consumption) and transfers in may contribute to the consumption of the recipient. Even though we do not know what the counterfactual consumptions would be in the absence of transfers, we deal with this to some extent in Table 17 by subtracting transfers from the recipient's consumption and adding transfers to the donor's consumption to build proxies for pre-transfer income.

Fourth, while we have information on ladder estimates for people living in the same location, we do not have information on their assets and therefore no Ω estimate. One way to deal with this would be to assume perfect information for these dyads and place their Ω values to 0. The results of this exercise are given in Table 18.

The results from these robustness checks are in line with our main results. The only difference is that in some specifications with either the Recipient Fixed Effect or the Donor Fixed Effect, the coefficient on Ω_{DR} becomes positive and weakly significant. However, our preferred specification to estimate the coefficients of Ω_{DR} and Ω_{RD} is the two-way fixed effects model of Equation 17, while our estimates for I_D and I_R come from the recipient's fixed effects model in Equation 18 and the donor's fixed effects model in Equation 19, respectively. Accordingly, all of the robustness checks yield coefficients from the preferred regressions that are consistent with the main results.

6.3 Robustness to discrete choice modeling

One shortcoming of the LPM model is that it makes extreme assumptions on the distributions of the error terms, which are likely to be violated in the case of a discrete outcome. To check for robustness we first use the conditional logit model of Chamberlain (1980), which provides consistent Fixed Effect estimators of the parameters $\beta_1, \beta_2, \beta_3$ and β_4 displayed in equations (18) and (19) when the outcome considered is discrete.

The results presented in columns (1) and (2) for Ω and in columns (5) and (6) for Ω' of Table 11 are qualitatively very similar to our main results in Table 9, which is reassuring.

In line with our models of pressure or exchange with recipient having full bargaining power, the results show that the higher the over-perception of income of donor the higher the transfers. Note that the very weak level of significance of Ω_{RD} in the regressions based on donor reports with Donor Fixed Effect (columns 2 and 6) may be explained by the particularly low number of observations contributing to the identification of the partial correlation.[12]

However, fixed effect (FE) type methods suffer from substantial efficiency losses as compared to methods based on the random effect (RE) principle and suffer from inconsistency biases if there are measurement errors. The latter may be concerning with respect to our main variables of interest Ω, $\Omega,'$, I_R and I_D.[13] Moreover, the double fixed effect approach provides no estimates for the donor and recipient incomes, which are important to distinguish the predictions of our theoretical models summarized in Table 8. Finally, they do not lend themselves easily to estimate discrete choice models. For all these reasons, one may be concerned about the precise identification of the main effects of interest $\beta_1, \beta_2, \beta_3$ and β_4 using FE-based methods displayed in equations (17) to (19).

To tackle these issues we estimated a Modified Random Effect model in line with Hajivassiliou (2012), which extends the Mundlak-Chamberlain approach by characterizing the correlations between the unobserved persistent heterogeneities α_i ($i = R$ in equation (18) or $i = D$ in equation

[12] A usual shortcoming of this approach is that the identification comes only from the switchers, which explains the low number of observations reported at the bottom of the Table.

[13] However, we are less worried about measurement errors concerning the variable Ω than concerning Ω', which is one of the reasons why we may prefer the Ω measure for our results.

Table 11: Discrete Choice Models, Partial correlations with transfers: Ω and income

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Ω_{RD}	0.391**	0.288	0.218***	0.231**				
	(2.24)	(1.39)	(2.76)	(2.13)				
Ω_{DR}	0.261	0.073	0.151	0.041				
	(1.46)	(0.44)	(1.57)	(0.46)				
Ω'_{RD}					0.891***	0.286	0.467***	0.433***
					(5.06)	(1.38)	(6.76)	(5.22)
Ω'_{DR}					0.245	-0.094	-0.025	-0.110
					(1.38)	(-0.56)	(-0.38)	(-1.38)
I_D	0.836***		0.440***	0.554***	1.514***		0.796***	0.881***
	(6.82)		(9.81)	(9.47)	(8.06)		(11.07)	(10.22)
I_R		-0.019	-0.049	-0.021		-0.102	-0.073	-0.116
		(-0.17)	(-1.27)	(-0.43)		(-0.64)	(-1.18)	(-1.43)
N	4197	3664	9032	9032	4197	3664	9032	9032

Notes: Discrete Choice Models with t values in parentheses under the coefficient. Estimated coefficients are shown * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Column 1 (2) presents results from Recipient (Donor) Fixed Effect Conditional Logit. Column 3 (4) presents results from a modified Donor (Recipient) Random Effect Probit model. Columns 5 to 8 replicate these methods using Ω' measures instead of Ω measures

(19)), and regressors as follows:

$$E(\alpha_i | \Omega_{RD}, \Omega_{DR}, \mathbf{P}_{RD}, I_D, I_R, \mathbf{Z}_D, \mathbf{Z}_R) \quad = \quad \mu_i = g_i(\Omega, P, I, Z)$$

assuming that $g_i(.)$ is a linear function of the regressors, that $g_i(.)$ depends only on the regressor data for individual i and that $g_i(.)$ only depends on the regressors in a household j invariant way.

This method involves simply adding the fixed household i regressors and family network averages of household j varying regressors as additional regressors in the right hand side of the models specified by equation (18) and equation (19) and proceeding with the RE Probit estimator to obtain consistent and efficient estimates. Note that we also nest the Random Effects within the extended family network clusters to account for possible correlations.

Columns (3)-(4) of Table 11 for Ω and columns (7)-(8) for Ω' show that our main results are robust to using a modified RE approach to control for the unobserved heterogeneities that may characterize Recipients or Donors of transfers. As with the FE approach, we find significant and positive correlations between transfers and misperceptions of donor's income but no significant correlations with misperceptions of recipient's income, once controlled for both donor and recipient actual incomes. We also find that the income of the donor is positively correlated to the transfers, as expected with a model of pressure to give or exchange with recipient having full bargaining power, and that the income of the receiver is clearly not significantly correlated to the transfers, rejecting the model under altruism or exchange with donor having full bargaining power.

7 Conclusion

This article addressed two main questions: what is the degree of misperceptions of income between households belonging to the same family network, and how do misperceptions of income relate to the private transfers between them? To do this we built novel measures of asymmetric information based on assets mutually perceived among households belonging to a same network. We apply these measures to original survey data collecting information in 2010 on households living in Tanzania, who are originated from the same families living in the Kagera area twenty years before the survey. Using the rich information available on their relationships and characteristics, we validated our measures and show that they capture well the asymmetry of information prevailing within these extended family networks. Interestingly we do not find that households are massively deceived in their perceptions of income, even when relatives have migrated to urban areas. As expected, the degree of misperceptions increases with genetic, social and physical distance.

We then developed simple static models to predict the correlations between income, misperceptions of income and transfers when altruism, exchange or pressure is the main driver of transfers. We show that the predictions of these models differ. Misperception is the the difference between perceived and actual income, such that positive values indicate over-estimation of income and negative values under-estimation. We show that in a model of pressure to give or in a model of exchange in which the recipient holds all the bargaining power, transfers have a positive partial correlation to the donor's actual income and the recipient's misperception of that income. In contrast, the income of the recipient and the misperception of it by the donor have a negative partial correlation to transfers under pure altruism. Finally, the recipient's income is negatively correlated to transfers under an exchange model in which the donor holds the bargaining power.

After examining such correlations in our data, we find evidence in support of a model of pressure to give or a model of exchange in which the recipient holds all bargaining power. These findings are important to improve the understanding of bilateral relationships and misperceptions of income existing between two linked households in an extended family, which has been rarely documented so far.

References

Altonji, Joseph G, Fumio Hayashi, and Laurence J Kotlikoff. 1997. "Parental Altruism and Inter Vivos Transfers: Theory and Evidence." *Journal of Political Economy* 105 (6):1121–66.

Ambler, Kate. 2013. "Don't Tell on Me: Experimental Evidence of Asymmetric Information in Transnational Households." Tech. rep., University of Michigan.

Attanasio, Orazio P. and Nicola Pavoni. 2011. "Risk Sharing in Private Information Models With Asset Accumulation: Explaining the Excess Smoothness of Consumption." *Econometrica* 79 (4):1027–1068.

Batista, Catia and Gaia Narciso. 2013. "Migrant Remittances and Information Flows: Evidence from a Field Experiment." CReAM Discussion Paper Series 1331, Centre for Research and Analysis of Migration (CReAM), Department of Economics, University College London.

Baum, Christopher, Mark Shaffer, and Steven Stillman. 2007. "Enhanced routines for instrumental variables/GMM estimation and testing." *The Stata Journal* 7 (4):465–506.

Becker, Gary S. 1974. "A Theory of Social Interactions." *Journal of Political Economy* 82 (6):pp. 1063–1093.

Beegle, Kathleen, Joachim De Weerdt, and Stefan Dercon. 2011. "Migration and Economic Mobility in Tanzania: Evidence from a Tracking Survey." *Review of Economics and Statistics* 93 (3):1010–1033.

Cameron, Collin, Jonah Gelbach, and Douglas Miller. 2011. "Robust Inference with Multi way Clustering." *Journal of Business Economics and Statistics* 29 (2):238–249.

Chort, Isabelle, Flore Gubert, and Jean-Noel Senne. 2012. "Migrant networks as a basis for social control: Remittance incentives among Senegalese in France and Italy." *Regional Science and Urban Economics* 42 (5):858 – 874. Special issue on Migration and Development.

Cole, Harold L. and Narayana R. Kocherlakota. 1999. "Efficient allocations with hidden income and hidden storage." Staff Report 238, Federal Reserve Bank of Minneapolis.

Cox, Donald, Zekeriya Eser, and Emmanuel Jimenez. 1998. "Motives for private transfers over the life cycle: An analytical framework and evidence for Peru." *Journal of Development Economics* 55 (1):57–80. URL `http://ideas.repec.org/a/eee/deveco/v55y1998i1p57-80.html`.

Cox, Donald and Marcel Fafchamps. 2008. *Extended Family and Kinship Networks: Economic Insights and Evolutionary Directions*, *Handbook of Development Economics*, vol. 4, chap. 58. Elsevier, 3711–3784.

De Weerdt, Joachim. 2004. *Risk-sharing and Endogenous Network Formation*, chap. 10. Handbook of Empirical Economics and Finance. Oxford University Press, 197–216.

———. 2010. “Moving out of Poverty in Tanzania: Evidence from Kagera.” *Journal of Development Studies* 46 (2):331–349.

Deaton, Angus. 1997. *The Analysis of Household Surveys: A Microeconometric Approach to Development Policy.* Baltimore Maryland: Johns Hopkins University Press.

Fafchamps, Marcel and Flore Gubert. 2007. “The formation of risk sharing networks.” *Journal of Development Economics* 83 (2):326–350.

Genicot, G. 2014. “Two-way Altruism and Signaling.” Tech. rep., Georgetown University.

Hajivassiliou, Vassilis. 2012. “Estimation and Specification Testing of Panel Data Models with Non-Ignorable Persistent Heterogeneity,Contemporaneous and Intertemporal Simultaneity, and Regime Classification Errors.” Tech. rep., London School of Economics.

Hauser, Christine and Hugo Hopenhayn. 2008. “Trading Favors: Optimal Exchange and Forgiveness.” Carlo Alberto Notebooks 88, Collegio Carlo Alberto. URL `http://ideas.repec.org/p/cca/wpaper/88.html`.

Jakiela, Pamela and Owen Ozier. 2012. “Does Africa Need a Rotten Kin Theorem? Experimental Evidence from Village Economies.” Tech. rep., World Bank Policy Research Working Paper 6085.

Kazianga, H. 2006. “Motives for household private transfers in Burkina Faso.” *Journal of Development Economics* 79 (1):73–117.

Lucas, Robert E B and Oded Stark. 1985. “Motivations to Remit: Evidence from Botswana.” *Journal of Political Economy* 93 (5):901–18.

McKenzie, David, John Gibson, and Steven Stillman. 2012. “A land of milk and honey with streets paved with gold: Do emigrants have over-optimistic expectations about incomes abroad?” *Journal of Development Economics* .

Mittag, Nikolas. 2012. "New methods to estimate models with large sets of fixed effects with an application to matched employer-employee data from Germany." FDZ Methodenreport 201201, Institut fur Arbeitsmarkt und Berufsforschung (IAB), Nurnberg [Institute for Employment Research, Nuremberg, Germany].

Platteau, Jean-Philippe. 2012. *Redistributive Pressures in Sub-Saharan Africa: Causes, Consequences, and Coping Strategies.* African Development in Historical Perspective, Cambridge University Press.

Rapoport, Hillel and Frederic Docquier. 2006. *The economics of migrants' remittances*, *Handbook on the Economics of Giving : Reciprocity and Altruism*, vol. 1, chap. 17. Elsevier, 1135–1198.

Serror, Marlon. 2012. "Measuring Information Asymmetries and Modeling their Impact on Senegalese Migrants Remittances." Tech. rep., Master's Thesis, Paris School of Economics and UC Berkeley.

Seshan, Ganesh. 2013. "Does Asymmetric Information Within Transnational Households Affect Remittance Flows?" Tech. rep., Georgetown University,- SFS Qatar.

Stark, Oded. 1995. *Altruism and Beyond: An Economic Analysis of Transfers and Exchanges within Families and Groups.* Cambridge: Cambridge University Press.

Appendix 1: Full set of recursive regressions leading to determination of weights in Ω

Table 12: Consumption Regressions

	First pass	**Second pass**	**Final regression**
HH located in urban area	0.204***	0.242***	0.243***
	(3.317)	(6.269)	(6.313)
Panel member finished O level	0.159*	0.176***	0.176***
	(1.892)	(5.883)	(5.874)
Panel member has a formal job	0.117	0.130**	0.177***
	(1.292)	(2.029)	(5.460)
Owns house	-0.116**	-0.079***	-0.083***
	(-2.536)	(-2.876)	(-3.686)
Owns land	-0.018	-0.014	
	(-0.261)	(-0.401)	
Owns livestock	-0.090	0.146***	0.146***
	(-0.567)	(4.920)	(4.924)
Owns phone	0.273***	0.308***	0.308***
	(4.279)	(14.135)	(14.191)
Owns TV, video equipment or camera	0.373***	0.314***	0.313***
	(4.995)	(12.006)	(11.987)
Owns motorbike, car, truck or other vehicle	0.210*	0.149*	0.147*
	(1.789)	(1.774)	(1.756)
Urban * completed O level	0.059		
	(0.952)		
Urban * has formal job	-0.003		
	(-0.043)		
Urban * owns house	0.023		
	(0.394)		
Urban * owns land	0.016		
	(0.210)		
Urban * owns livestock	-0.138**	-0.137**	-0.136**
	(-2.134)	(-2.152)	(-2.129)
Urban * owns phone	-0.099**	-0.114**	-0.115***
	(-2.044)	(-2.561)	(-2.579)
Urban * owns TV	-0.043		
	(-0.778)		
Urban * owns vehicle	0.132**	0.139**	0.141**

Continued on next page...

	First pass	Second pass	Final regression
	(2.013)	(2.260)	(2.284)
Owns house * completed O level	-0.006		
	(-0.073)		
Owns house * has formal job	-0.130		
	(-1.373)		
Owns house * owns livestock	0.008		
	(0.088)		
Owns house * owns phone	0.032		
	(0.508)		
Owns house * owns TV	0.074		
	(1.016)		
Owns house * owns vehicle	0.257**	0.235***	0.237***
	(2.251)	(2.999)	(3.026)
Owns land * completed O level	-0.015		
	(-0.154)		
Owns land * has formal job	0.183*	0.060	
	(1.648)	(0.846)	
Owns land * owns phone	0.006		
	(0.072)		
Owns land * owns vehicle	-0.089		
	(-0.595)		
Owns land * owns livestock	0.235		
	(1.282)		
Owns land * owns TV	-0.097		
	(-1.131)		
Head is male	0.011	0.009	0.008
	(0.396)	(0.304)	(0.300)
Age of hh head	-0.007*	-0.007*	-0.007*
	(-1.851)	(-1.925)	(-1.900)
Head age squared	0.000	0.000	0.000
	(1.112)	(1.161)	(1.135)
Educ of hh head (in years)	0.018***	0.019***	0.019***
	(5.473)	(5.637)	(5.615)
Males 0-5 years	-0.219***	-0.219***	-0.219***
	(-17.901)	(-17.946)	(-17.957)
Males 6-15 years	-0.105***	-0.105***	-0.105***
	(-10.390)	(-10.431)	(-10.464)
Males 16-60 years	-0.000	0.000	0.000
	(-0.033)	(0.015)	(0.024)

Continued on next page...

	First pass	Second pass	Final regression
Males 61+ years	0.038	0.040	0.040
	(0.867)	(0.911)	(0.908)
Females 0-5 years	-0.206***	-0.207***	-0.207***
	(-16.641)	(-16.777)	(-16.802)
Females 6-15 years	-0.101***	-0.101***	-0.101***
	(-9.881)	(-9.935)	(-9.949)
Females 16-60 years	-0.048***	-0.048***	-0.048***
	(-3.697)	(-3.728)	(-3.723)
Females 61+ years	-0.090***	-0.091***	-0.091***
	(-2.687)	(-2.707)	(-2.718)
Constant	13.301***	13.272***	13.260***
	(140.777)	(162.610)	(165.682)
Adjusted R-squared	0.575	0.576	0.576
N	3173	3173	3173

Notes: Recursive regressions leading to estimation of Equation (3). t statistics in brackets under the coefficient. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Appendix 2: Proofs

Model of Pressure

We can rewrite the incentive constraint for type L in (10) as

$$u_D(L) - u_D(L - T_L) \leq P - \underline{p}. \tag{23}$$

It follows then directly from the concavity of u_D that

$$u_D(H) - u_D(H - T_L) \leq P - \underline{p},$$

so that, if (23) holds a donor of type H would prefer pretending to be a low type than make zero transfer and receive full pressure. It follows that the relevant constraint for type H is not to want to pretend to be a low type (11). We can rewrite this constraint as

$$u_D(H - T_L) - u_D(H - T_H) \leq \underline{p}. \tag{24}$$

Given his income x and the signal s received, the recipient R chooses $\underline{p}$, T_H and T_L to maximize

$$\pi_s^D u_R(x + T_H) + (1 - \pi_s^D) u_R(x + T_L) \tag{25}$$

subject to (23) and (24).

Denoting as λ and μ the Lagrange multipliers on (23) and (24) respectively, the first order conditions tell us that

$$\pi_s^D u_R'(x + T_H) = \mu u_D'(H - T_H) \tag{26}$$

$$(1 - \pi_s^D) u_R'(x + T_L) = \lambda u_D'(L - T_L) - \mu u_D'(H - T_L), \tag{27}$$

and that $\lambda = \mu$ if $\underline{p} \in (0, P)$, while $\underline{p} = 0$ if $\mu < \lambda$ and $\underline{p} = P$ if $\mu > \lambda$.

Proof of Proposition 1

Assumption S1 implies that π_s^D is increasing in s, while Assumption S2 implies that π_s^D tends to 0 when s tends to 0 and π_s^D tends to 1 when s tends to 1.

For values of the signal s close to 0, π_s^D is close to 0. Equations (26) and (27) imply then that μ is close to 0 while λ is strictly positive. It follows that $\underline{p}$ must be 0 and $T_L = T_H$. The recipient offers a pooling contract. It is obvious from the objective that higher values of the signal gives incentive to raise T_H and to lower T_L, thereby making a separating contract more likely. When the signal s takes values close to 1, $(1 - \pi_s^D)$ and therefore the left hand side of (27) is close to 0. Since $u'_D(L - T_L) > u'_D(H - T_L)$ for any $T_L > 0$, it must be that $\mu > \lambda$. Hence, $\underline{p} = P$ and $T_L = 0$.

Thus, there is a cutoff value of the signal $\widetilde{s} \in (0, 1)$ that is such that R offers a pooling contract if $s_D \leq \widetilde{s}$ and R offers a separating contract if $s_D > \widetilde{s}$ with T_H (T_L) increasing (decreasing) in s.

Appendix 3: Summary statistics

Table 13: Summary statistics of household variables

	Mean	**SD**
Log consumption per capita	13.12	0.70
HH located in urban area	0.34	0.47
Panel member finished O level	0.13	0.33
Panel member has a formal job	0.09	0.28
Owns house	0.75	0.43
Owns land	0.87	0.34
Owns livestock	0.12	0.33
Owns phone	0.60	0.49
Owns TV, video equipment or camera	0.19	0.39
Owns motorbike, car, truck or other vehicle	0.09	0.29
Head is male	0.80	0.40
Age of hh head	41.04	15.14
Educ of hh head (in years)	6.35	3.26
Males 0-5 years	0.50	0.71
Males 6-15 years	0.60	0.87
Males 16-60 years	1.07	0.73
Males 61+ years	0.08	0.27
Females 0-5 years	0.49	0.70
Females 6-15 years	0.61	0.86
Females 16-60 years	1.11	0.74
Females 61+ years	0.11	0.33

Notes: $N = 3,173$.

Table 14: Summary statistics of dyadic variables

	Sample $N = 9,032$		Sample $N = 13,808$	
	Mean	**SD**	**Mean**	**SD**
D reports giving gift to R in past 12 months	0.20	0.40	0.24	0.42
Ω_{RD}	0.01	0.29	0.01	0.29
Ω_{DR}	0.01	0.29	0.01	0.29
Ω'_{RD}	0.01	0.52	0.00	0.52
Ω'_{DR}	0.01	0.52	0.00	0.52
R places D low on ladder ($\underline{L}_{RD}$)	0.41	0.49	0.42	0.49
R places D high on ladder ($\overline{L}_{RD}$)	0.05	0.21	0.04	0.21
R answers DK on ladder question about D	0.11	0.31	0.09	0.28
D places R low on ladder ($\underline{L}_{DR}$)	0.41	0.49	0.42	0.49
D places R high on ladder ($\overline{L}_{DR}$)	0.05	0.21	0.04	0.21
D answers DK on ladder question about R	0.11	0.31	0.09	0.28
I_R	13.19	0.72	13.13	0.69
R HH head is male	0.82	0.39	0.81	0.39
R HH head age	39.93	14.22	40.51	14.77
R HH head years education	6.74	3.27	6.40	3.23
R HH No. Males 0-5 years	0.49	0.69	0.51	0.71
R HH No. Males 6-15 years	0.60	0.87	0.62	0.89
R HH No. Males 16-60 years	1.08	0.74	1.07	0.71
R HH No. Males 61+ years	0.06	0.25	0.07	0.26
R HH No. Females 0-5 years	0.49	0.69	0.50	0.70
R HH No. Females 6-15 years	0.60	0.86	0.62	0.87
R HH No. Females 16-60 years	1.12	0.74	1.12	0.75
R HH No. Females 61+ years	0.10	0.31	0.11	0.32
Parent-child link	0.17	0.38	0.22	0.41
km distance	214.07	339.40	150.51	295.71

Appendix 4: Robustness Tests

Table 15: Table 9 using transfers in

	Equation					
	(17)	**(18)**	**(19)**	**(17)**	**(18)**	**(19)**
Ω_{RD}	0.075*	0.039**	0.133***			
	(1.84)	(2.21)	(5.45)			
Ω_{DR}	0.048	0.039*	0.042**			
	(1.22)	(1.76)	(2.13)			
Ω'_{RD}				0.075*	0.129***	0.132***
				(1.84)	(6.91)	(5.42)
Ω'_{DR}				0.048	0.039*	0.036*
				(1.22)	(1.78)	(1.75)
I_D		0.058***			0.158***	
		(5.56)			(8.49)	
I_R			0.014			0.038*
			(1.15)			(1.84)
N	9032	9032	9032	9032	9032	9032

Notes: Fixed Effect Linear probability models with t values in parentheses under the coefficient. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 16: Table 9, clustering at network level

	Equation					
	(17)	**(18)**	**(19)**	**(17)**	**(18)**	**(19)**
Ω_{RD}	0.093**	0.047**	0.056***			
	(2.12)	(2.30)	(2.71)			
Ω_{DR}	0.003	0.038*	0.009			
	(0.08)	(1.70)	(0.57)			
Ω'_{RD}				0.093**	0.101***	0.056***
				(2.12)	(4.66)	(2.69)
Ω'_{DR}				0.003	0.037*	0.007
				(0.08)	(1.69)	(0.40)
I_D		0.090***			0.167***	
		(6.20)			(7.25)	
I_R			0.007			0.011
			(0.67)			(0.72)
N	9032	9032	9032	9032	9032	9032

Notes: Fixed Effect Linear probability models with t values in parentheses under the coefficient. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 17: Table 9 using consumption purged of transfers

	Equation					
	(17)	**(18)**	**(19)**	**(17)**	**(18)**	**(19)**
Ω_{RD}	0.093**	0.049***	0.057***			
	(2.24)	(2.60)	(2.75)			
Ω_{DR}	0.003	0.037	0.007			
	(0.08)	(1.59)	(0.44)			
Ω'_{RD}				0.093**	0.114***	0.056***
				(2.24)	(5.45)	(2.73)
Ω'_{DR}				0.003	0.037	-0.008
				(0.08)	(1.58)	(-0.43)
I_D		0.095***			0.182***	
		(7.95)			(8.96)	
I_R			0.002			-0.005
			(0.16)			(-0.32)
N	9032	9032	9032	9032	9032	9032

Notes: Fixed Effect Linear probability models with t values in parentheses under the coefficient. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 18: Table 9 including dyads living in same location

	Equation					
	(17)	**(18)**	**(19)**	**(17)**	**(18)**	**(19)**
Ω_{RD}	0.093**	0.047**	0.056***			
	(2.24)	(2.49)	(2.74)			
Ω_{DR}	0.003	0.038	0.009			
	(0.08)	(1.60)	(0.58)			
Ω'_{RD}				0.093**	0.101***	0.056***
				(2.24)	(4.85)	(2.72)
Ω'_{DR}				0.003	0.037	0.007
				(0.08)	(1.59)	(0.38)
I_D		0.090***			0.167***	
		(7.54)			(8.22)	
I_R			0.007			0.011
			(0.73)			(0.66)
N	13808	13808	13808	13808	13808	13808

Notes: Fixed Effect Linear probability models with t values in parentheses under the coefficient. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Ω is set to 0 for dyads living in same location.